

The listing of claims will replace all prior versions, and listings, of claims in the application:

Listing of Claims:

B1
1. (Currently Amended) A method of providing load balancing among host servers in a computer network using a load balancing switch and a plurality of site switches, the method comprising:

coupling said load balancing switch between said computer network and an authoritative domain name system server and configuring said load balancing switch as a proxy to said authoritative domain name server;

coupling each of said host servers to said computer network through said site switches;

collecting at said load ~~balance~~ balancing switch a first set of performance metrics regarding said network;

whenever said authoritative domain name system server provides network addresses in response to a query regarding a domain name, arranging, at said load balancing switch, said network addresses as an ordered list in accordance with said performance metrics;

forwarding said ordered list of network addresses as a response to said query to an originator of said query.

2. (Original) A method as in Claim 1, further comprising:

collecting a second set of performance metrics regarding said network, said second set of performance metrics reflecting access conditions to said host servers at each of said site switches;

B1 sending said second set of performance metrics from said site switches to said load balancing server; and

including said second set of performance metrics with said first set of performance metrics.

3. (Original) The method of Claim 1, wherein said first set of performance metrics includes a health check sent from said load balancing switch to each of said site switches.

4. (Currently Amended) The method of Claim 3 wherein, when a particular site switch ~~any of said host servers~~ fails said health check, a network address of said ~~failed host server~~ particular site switch is provided a lesser position in said ordered list.

5. (Currently Amended) The method of Claim 2, wherein said ~~collection~~ collecting of said second set of performance metrics includes recording, at each site switch, a number of sessions connected to host servers having network addresses configured on said ~~said~~ the respective site switch.

6. (Currently Amended) The method of Claim 5, wherein when said number of sessions at said a particular site switch exceeds a ~~predetermined~~ threshold percentage of ~~that~~ said

particular site switch's maximum capacity, a corresponding one of said network addresses is provided a lesser position in said ordered list.

B1

7. (Currently Amended) The method of Claim 2, wherein said collecting said second set of performance metrics includes recording, at each site switch, a round trip time indicative of elapse time for exchanging messages between ~~each~~ the respective site switch and a client machine of said computer network.

8. (Currently Amended) The method of Claim 7, wherein said round trip time ~~being~~ is an actual recorded time period between said the respective site switch receiving a connection request from said client machine and said the respective site switch receiving an acknowledgement of a connection from said client machine.

9. (Original) The method of Claim 1, wherein said arranging takes into consideration the geographical location of said originator of said query.

10. (Original) The method of Claim 3, wherein said collecting of said first set of performance metrics includes recording a time interval for each site switch between said load balancing switch initiating said health check and said load balancing switch receiving a response from said site switch.

B1
11. (Original) The method of Claim 1, wherein said arranging selects a network address of a least recently selected host server for placement at a higher position in said ordered list.

12. (Original) The method of Claim 1 further comprising said load balancing switch limiting a valid time for each network address in said ordered list to less than a predetermined value.

13. (Currently Amended) The method of Claim 1, further comprising, when a connection request is received at a particular site switch for a connection to one of said host servers, said particular site switch redirecting said connection request to another one of said host servers.

14. (Currently Amended) A system for balancing load among host servers in a computer network, comprising:

an authoritative domain name system server;

a load balancing switch coupled to said authoritative domain name system server, said load balancing switch (a) being configured to be a proxy to said authoritative domain name system server; (b) collecting a first set of performance metrics regarding said network; and (c) arranging a list of network addresses from said authoritative domain name system server in accordance with said first set of performance metrics; and

a plurality of site switches coupling said host servers to said computer network.

B1

15. (Currently Amended) A system as in Claim 14, wherein each of said site switches (a) collects a second set of performance metrics regarding said network, said second set of performance metrics reflecting access conditions to host servers at said the respective site switch; and (b) sends said second set of performance metrics to said load balancing switch server; whereupon said load balancing switch includes said second set of performance metrics with said first set of performance metrics.

16. (Original) A system as in Claim 14, wherein said first set of performance metrics includes a health check sent from said load balancing switch to each of said site switches.

17. (Currently Amended) A system as in Claim 16 wherein, when ~~any of said~~ a particular ~~host servers~~ site switch fails said health check, a network address of said ~~failed~~ particular site switch ~~host server~~ is provided a lesser position in said ordered list.

18. (Currently Amended) A system as in Claim 15, wherein said second set of performance metrics includes a number of sessions connected to a network address configured at said respective site switch.

B1 19. (Currently Amended) A system as in Claim 18, wherein when said number of sessions exceeds a ~~predetermined~~ threshold percentage of ~~that~~ said respective site switch's maximum capacity, said network address is provided a lesser position in said ordered list.

20. (Currently Amended) A system as in Claim 15, wherein said second set of performance metrics includes a round trip time indicative of elapse time for exchanging messages between each host server coupled to said respective site switch and a client machine of said computer network.

21. (Currently Amended) A system as in Claim 20, wherein said round trip time ~~being~~ is an actual recorded time period between said respective site switch receiving a connection request from said client machine and said respective site switch receiving an acknowledgement of a connection from said client machine.

22. (Original) A system as in Claim 14, wherein said arranging takes into consideration the geographical location of said originator of said query.

23. (Currently Amended) A system as in Claim 17, wherein said first set of performance metrics includes a time interval for each site switch between said load balancing

switch initiating said health check and said load balancing switch receiving a response from ~~said~~ the
respective site switch.

B1

24. (Original) A system as in Claim 14, wherein said arranging selects a network address of
a least recently selected host server for placement at a higher position in said ordered list.

25. (Original) A system as in Claim 14, wherein said load balancing switch limits a valid
time for each network address in said ordered list to less than a predetermined value.

26. (Original) A system as in Claim 14, wherein when a connection request is received at a
site switch for a connection to one of said host servers, said site switch redirects said connection
request to another one of said host servers.

27. (Currently Amended) A method for providing load balancing among a plurality
of host servers in a computer network, the method comprising:

receiving a request at a load balancing switch from a client relating to ~~any~~ one of
the plurality of host servers;

forwarding the request to an authoritative domain name system server distinct
from the load balancing switch;

identifying at the authoritative domain name system server one or more of the plurality of host servers;

B1 ordering the one or more host servers at the load balancing switch based on performance metrics including at least a round trip time associated with the client; and

sending a response to the client including information associated with each of the one or more identified host servers.

28. (Currently Amended) The method of claim 27, wherein access to each of the plurality of host servers is controlled by one of a plurality of computers, and

wherein the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers.

29. (Currently Amended) The method of claim 28, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include an indication of whether the number of sessions communicating through the respective computer exceeds a predetermined threshold.

30. (Original) The method of claim 27, wherein the receiving comprises receiving a request from the client to resolve a domain name associated with any one of the plurality of host servers.

B1
31. (Currently Amended) The method of claim 27, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics including an indication of the health of one or more of the plurality of host servers.

32. (Currently Amended) The method of claim 27, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics including an indication of the health of an application on one or more of the plurality of host servers.

33. (Cancelled)

34. (Currently Amended) The method of claim 27, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics including a geographical location associated with one or more of the plurality of host servers.

35. (Currently Amended) The method of claim 28, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include an indication of the available session capacity of the respective computer.

B1 36. (Currently Amended) The method of claim 28, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include a time required by the respective computer to provide an indication of the health of a host server access to which is controlled by the respective computer.

37. (Currently Amended) The method of claim 28, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host computers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include a time required by the respective computer to provide an indication of the health of an application on a host server access to which is controlled by the respective computer.

38. (Currently Amended) The method of claim 27, wherein the ordering ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on the number of times each of the one or more plurality of host servers has been previously identified.

39. (Original) The method of claim 27, wherein the sending comprises sending a response to the client including information associated with each of the one or more identified

servers, wherein the information includes one or more network address each of which is associated with one of the one or more identified host servers.

B1

40. (Currently Amended) The method of claim 39, wherein the sending comprises sending a response to the client including information associated with each of the one or more identified servers, wherein the information includes one or more network address each of which is associated with one of the one or more identified host servers, ~~and wherein the one or more network addresses are ordered based on the performance metrics.~~

41. (Cancelled)

42. (Currently Amended) A system for providing load balancing among a plurality of host servers in a computer network, the system comprising:

means for receiving a request from a client relating to ~~any~~ one of the plurality of host servers;

means for forwarding the request to a means for identifying one or more of the plurality of host servers;

means for ordering, distinct from the means for identifying, the one or more host servers based on performance metrics including at least a round trip time associated with the client; and

means for sending a response to the client including information associated with each of the one or more identified host servers.

43. (Currently Amended) The system of claim 42, wherein access to each of the plurality of host servers is controlled by one of a plurality of computers, and

B1 wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics collected from each of the plurality of computers.

44. (Currently Amended) The system of claim 43, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include an indication of whether the number of sessions communicating through the respective computer exceeds a predetermined threshold.

45. (Original) The system of claim 42, wherein the means for receiving comprises means for receiving a request from the client to resolve a domain name associated with any one of the plurality of host servers.

46. (Currently Amended) The system of claim 42, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics including an indication of the health of one or more of the plurality of host servers.

B1
47. (Currently Amended) The system of claim 42, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics including an indication of the health of an application on one or more of the plurality of host servers.

48. (Cancelled)

49. (Currently Amended) The system of claim 42, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics including a geographical location associated with one or more of the plurality of host servers.

50. (Currently Amended) The system of claim 43, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include an indication of the available session capacity of the respective computer.

51. (Currently Amended) The system of claim 43, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include a time

required by the respective computer to provide an indication of the health of a host server access to which is controlled by the respective computer.

B1

52. (Currently Amended) The system of claim 43, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host computers based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include a time required by the respective computer to provide an indication of the health of an application on a host server access to which is controlled by the respective computer.

53. (Currently Amended) The system of claim 42, wherein the means for ~~identifying~~ ordering comprises means for ~~identifying~~ ordering one or more of the plurality of host servers based on the number of times each of the one or more plurality of host servers has been previously identified.

54. (Original) The system of claim 42, wherein the means for sending comprises means for sending a response to the client including information associated with each of the one or more identified servers, wherein the information includes one or more network address each of which is associated with one of the one or more identified host servers.

55. (Original) The system of claim 54, wherein the means for sending comprises means for sending a response to the client including information associated with each of the one or more identified servers, wherein the information includes one or more network address each

of which is associated with one of the one or more identified host servers, and wherein the one or more network addresses are ordered based on the performance metrics.

B1

56. (Currently Amended) A computer program product comprising a computer readable medium having computer readable code embodied therein, the computer readable code, when executed, causing a computer to implement a method for providing load balancing among a plurality of host servers in a computer network, the method comprising:

receiving a request from a client relating to ~~any~~ one of the plurality of host servers;

forwarding the request to an authoritative domain name system server;

receiving a response from the authoritative domain name system server

identifying one or more of the plurality of host servers;

ordering the one or more host servers based on performance metrics including at least a round trip time associated with the client; and

sending a response to the client including information associated with each of the one or more identified host servers.

57. (Currently Amended) The computer program product of claim 56, wherein access to each of the plurality of host servers is controlled by one of a plurality of computers, and wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers.

B1

58. (Currently Amended) The computer program product of claim 57, wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include an indication of whether the number of sessions communicating through the respective computer exceeds a predetermined threshold.

59. (Original) The computer program product of claim 56, wherein, in the implemented method, the receiving comprises receiving a request from the client to resolve a domain name associated with any one of the plurality of host servers.

60. (Currently Amended) The computer program product of claim 56, wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host server~~ ordering based on performance metrics including an indication of the health of one or more of the plurality of host servers.

61. (Currently Amended) The computer program product of claim 56, wherein, in the implemented method, the ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics including an indication of the health of an application on one or more of the plurality of host servers.

62. (Cancelled)

B1 63. (Currently Amended) The computer program product of claim 56, wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics including a geographical location associated with one or more of the plurality of host servers.

64. (Currently Amended) The computer program product of claim 57, wherein, in the implemented method, the ~~identifying~~ comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include an indication of the available session capacity of the respective computer.

65. (Currently Amended) The computer program product of claim 57, wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host servers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include a time required by the respective computer to provide an indication of the health of a host server access to which is controlled by the respective computer.

66. (Currently Amended) The computer program product of claim 57, wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host computers~~ ordering based on performance metrics collected from each of the plurality of computers, wherein the performance metrics collected from each of the plurality of computers include a time required by the respective computer to provide an indication of the

health of an application on a host server access to which is controlled by the respective computer.

B1

67. (Currently Amended) The computer program product of claim 56, wherein, in the implemented method, the ~~identifying~~ ordering comprises ~~identifying one or more of the plurality of host servers~~ ordering based on the number of times each of the one or more plurality of host servers has been previously identified.

68. (Original) The computer program product of claim 56, wherein, in the implemented method, the sending comprises sending a response to the client including information associated with each of the one or more identified servers, wherein the information includes one or more network address each of which is associated with one of the one or more identified host servers.

69. (Original) The computer program product of claim 68, wherein, in the implemented method, the sending comprises sending a response to the client including information associated with each of the one or more identified servers, wherein the information includes one or more network address each of which is associated with one of the one or more identified host servers, and wherein the one or more network addresses are ordered based on the performance metrics.